THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 28, Issue of July 14, pp. 21099–21106, 2000
Printed in U.S.A.
© 2000 by The American Society for Biochemistry and Molecular Biology, Inc.

# Characterization of Alzheimer's β-Secretase Protein BACE

## A PEPSIN FAMILY MEMBER WITH UNUSUAL PROPERTIES*

Received for publication, March 13, 2000, and in revised form, April 24, 2000
Published, JBC Papers in Press, April 28, 2000, DOI 10.1074/jbc.M002095200

**Mitsuru Haniu‡, Paul Denis, Yunjen Young, Elizabeth A. Mendiaz, Janis Fuller, John O. Hui, Brian D. Bennett, Steven Kahn, Sandra Ross, Teresa Burgess, Viswanatham Katta, Gary Rogers, Robert Vassar, and Martin Citron**

*From Amgen Inc., Thousand Oaks, California 91320-1799*

**The cerebral deposition of amyloid β-peptide is an early and critical feature of Alzheimer's disease. Amyloid β-peptide is released from the amyloid precursor protein by the sequential action of two proteases, β-secretase and γ-secretase, and these proteases are prime targets for therapeutic intervention. We have recently cloned a novel aspartic protease, BACE, with all the known properties of β-secretase. Here we demonstrate that BACE is an *N*-glycosylated integral membrane protein that undergoes constitutive N-terminal processing in the Golgi apparatus. We have used a secreted Fc fusion-form of BACE (BACE-IgG) that contains the entire ectodomain for a detailed analysis of post-translational modifications. This molecule starts at Glu$^{46}$ and contains four *N*-glycosylation sites (Asn$^{153}$, Asn$^{172}$, Asn$^{223}$, and Asn$^{354}$). The six Cys residues in the ectodomain form three intramolecular disulfide linkages (Cys$^{216}$–Cys$^{420}$, Cys$^{278}$–Cys$^{443}$, and Cys$^{330}$–Cys$^{380}$). Despite the conservation of the active site residues and the 30–37% amino acid homology with known aspartic proteases, the disulfide motif is fundamentally different from that of other aspartic proteases. This difference may affect the substrate specificity of the enzyme. Taken together, both the presence of a transmembrane domain and the unusual disulfide bond structure lead us to conclude that BACE is an atypical pepsin family member.**

The hallmarks of Alzheimer's disease (AD)[1] pathology are brain plaques and vascular deposits (1) consisting of the 4-kDa amyloid β-peptide (Aβ) (2). Overproduction of the 42-amino acid form of Aβ, Aβ42, has been suggested to be the cause of all known cases of familial early onset AD (3), and it is assumed that Aβ42 deposition plays an early and critical role in sporadic AD as well. Therefore, Aβ metabolism has attracted considerable interest. In 1987 it was shown (4) that formation of Aβ requires proteolytic cleavage of a large type I transmembrane protein, the β-amyloid precursor protein (APP), which is constitutively expressed in most cell types. Over the next decade the proteolytic processing of APP has been studied in great detail in a variety of systems by many groups. Taken together, these studies have shown that Aβ is generated at a low rate by most cells analyzed and that two different proteolytic activities are required for Aβ generation. First, β-secretase cleaves APP to generate the N terminus of Aβ, and second, γ-secretase cleaves the C terminus, leading to the release of Aβ (for review see Ref. 5). Studies with intact cells expressing APP and the endogenous secretases have led to conclusions about the properties of the β- and γ-secretases, *e.g.* their tissue distribution, subcellular localization, substrate requirements (see *e.g.* Ref. 6) etc., but until recently the identity of both β- and γ-secretase was unknown. This changed when we very recently identified the novel transmembrane aspartic protease BACE as the major β-secretase (7). Three subsequently published independent studies (8–10) have confirmed this conclusion. Here we characterize the BACE protein. We show that BACE is an *N*-glycosylated integral membrane protein that undergoes constitutive N-terminal processing in the Golgi apparatus. We determine the processing and *N*-glycosylation sites and the disulfide bonds. Our results demonstrate that BACE is an unusual member of the pepsin family.

## EXPERIMENTAL PROCEDURES

*Materials*—Trypsin, pepsin, and endoproteinase Asp-N were obtained from Roche Molecular Biochemicals. Fluorescein 5-maleimide (FM) was purchased from Molecular Probes (Eugene, OR). 4-HCCA was from Sigma. Sialidase was obtained from Glyko (Novato, CA). *N*- and *O*-glycanases were from Genzyme (Cambridge, MA). *N*-Glycosidase F was purchased from Roche Molecular Biochemicals. Other chemicals are of high quality grade.

*Analysis of BACE Membrane Binding*—Untransfected 293 cells or 293 cells stably expressing BACE were scraped into phosphate-buffered saline, and the cells were precipitated. The pellet was resuspended in 25 mM HEPES, pH 7.2, with protease inhibitors, and the cells were swollen on ice for 60 min. Cells were lysed by 3 freeze-thaw cycles at −80 °C and then centrifuged for 15 min at 1,000 × *g* to precipitate nuclei. The supernatant was centrifuged for 60 min at 100,000 × *g* to give a crude membrane pellet and a supernatant containing cytosolic proteins. Membranes were solubilized in 25 mM HEPES, pH 7.2, 2% CHAPS and centrifuged at 20,000 × *g* for 10 min. The resulting supernatant contained the membrane-bound proteins. To determine if BACE is an integral or peripheral membrane protein, crude membranes were washed with either 0.5 M NaCl or 100 mM $Na_2CO_3$, pH 11, to release peripherally bound proteins.

*Analysis of BACE Posttranslational Modifications*—A polyclonal antibody specific to the propeptide region of BACE was raised following standard procedures using as immunogen the peptide CGIRLPLRSGLGGAPLGLRLPR (comprising amino acids 25–45 of BACE and an N-terminal Cys residue for coupling). After metabolic labeling with [$^{35}$S]methionine aliquots of the same cell lysates were immunoprecipitated using the previously described BACE C-terminal antiserum (7) and the propeptide antiserum following protocols described before (11).

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Amgen Inc., One Amgen Center Dr., Thousand Oaks, CA 91320-1799. Tel.: 805-447-3117; Fax: 805-499-7464; E-mail: mhaniu@amgen.com.

[1] The abbreviations used are: AD, Alzheimer's disease; APP, amyloid precursor protein; BACE, β-site APP-cleaving enzyme; FM, fluorescein 5-maleimide; 4-HCCA, α-cyano-4-hydroxycinnamic acid; HPLC, high performance liquid chromatography; MALDI, matrix-assisted laser desorption ionization; TD-, trypsin-endoproteinase Asp-N; CHAPS, (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; Aβ, amyloid β-peptide; PAGE, polyacrylamide gel electrophoresis; PTH, phenylthiohydantoin.

*N*-Glycosidase F treatment was performed after immunoprecipitation. For pulse-chase experiments cells were metabolically labeled for 20 min and then chased for the indicated times. Brefeldin A, dissolved as a 30 mM stock in methanol, was used at 30 μM final concentration in medium during a 3-h chase. Immunoprecipitates were analyzed by SDS-PAGE followed by quantitative imaging on a STORM 860 phosphorimaging system (Molecular Dynamics).

*Preparation and Purification of BACE-IgG*—The BACE-IgG construct containing cDNA encoding the ectodomain of BACE (residues 1–460) and the Fc portion of human IgG1 (230 amino acids) was described previously (7). BACE-IgG protein was purified from conditioned media of stably transfected 293T cells with protein A columns. The protein A eluate consisted of BACE-IgG and a low level of clipped Fc fragment. In order to remove the Fc contaminant, this material was further purified by gel filtration using a Sephacryl S-300HR (Amersham Pharmacia Biotech) column (3.2 × 46 cm) in phosphate-buffered saline buffer.

*Treatment of the Enzyme with Fluorescein 5-Maleimide*—In order to examine the existence of free sulfhydryl residues in BACE-IgG, the sample was treated with 10 mM FM in 50 mM Tris-HCl, 4 M guanidine HCl, pH 7.5, at room temperature for 20 h. Excess reagents were removed by passing through reversed phase HPLC using a Vydac C18 column (2.1 × 150 mm). The protein fraction was subjected to proteolytic digestion for peptide mapping.

*Proteolytic Fragmentation of BACE-IgG*—The above FM-modified sample and intact BACE-IgG (~50 μg) were initially digested with trypsin (1 μg) at 37 °C for 20 h in 0.1 M Tris buffer, pH 7.5 (200 μl). The sample was allowed to proceed to a second digestion with endoproteinase Asp-N (1 μg) under the same conditions. The digested materials were directly subjected to reversed phase HPLC using a Vydac C18 column (2.1 × 150 mm). Peptic digestion of the protein (~50 μg) was performed in 0.02 N HCl, pH 2 (200 μl), for 20 h at 37 °C with an enzyme:substrate ratio of 1:50 (w/w), and the digestion was terminated by direct injection onto reversed phase HPLC.

*HPLC Separation of the Peptides*—Trypsin-endoproteinase-Asp-N (TD)- or pepsin (P)-generated peptides were separated by reversed phase HPLC using a Vydac C18 column (2.1 × 150 mm). Two solvent systems (solvents A and B) were utilized, solvent A (0.1% trifluoroacetic acid) and solvent B (0.1% trifluoroacetic acid, 90% acetonitrile). The peptides were eluted with a linear gradient from 2% solvent B to 40% solvent B over 40 min and second gradient from 40% solvent B to 60% solvent B over 10 min. Flow rate was constant at 0.25 ml/min. The peptide was detected by absorbances at 215 and 280 nm.

*Treatment with N- and O-Glycanases and Sialidase*—Glycoprotein or glycopeptides were treated with several enzymes. For removal of sialic acid, the dried protein sample was dissolved in 20 mM sodium acetate buffer, pH 5 (50 μl), and incubated with sialidase (0.1 unit) for 20 h at 37 °C. Protein samples were deglycosylated with *N*- and *O*-glycanases in 20 mM sodium acetate buffer, pH 5, and were subjected to SDS-PAGE. Glycopeptides were incubated with the above enzymes under the same conditions. The sample was purified by reversed phase HPLC for mass spectrometry.

*Mass Spectrometry of Disulfide Peptides and Glycopeptides*—Matrix-assisted laser desorption ionization (MALDI)-mass spectrometry of the peptides was performed using either a Kratos IV (Kratos Analytical) or Voyager mass spectrometer (PerSeptive Biosystems). The sample was dissolved in 0.1% trifluoroacetic acid, 50% acetonitrile and then spotted on the sample plate with sinapinic acid or 4-HCCA as matrix. Cys-containing peptides were also analyzed using an ion-spray interface using a Michrome BIOSOURCE Ultrafast Microprotein Analyzer. The carrier solvent was 50% acetonitrile:water with 0.1% trifluoroacetic acid flowing at 5 μl/min. The scan range was 300–2400 atomic mass units with a step of 0.5 atomic unit. The mass units and standard deviation were calculated using Sciex hypermass software.

*Amino Acid Sequence Analysis*—N-Terminal sequence analysis of peptides and proteins was performed on a model 494 ABI Procise sequencer system from Perkin-Elmer/Applied Biosystems Inc. (Foster City, CA). For analysis of PTH amino acids, an ABI 140 system was used. Data analysis was performed with the Applied Biosystems model 610 data analysis program for protein sequencing, version 2.1.

*Carbohydrate Analysis*—*N*-Glycosylation sites of the enzyme were identified by negative response on PTH analysis at the corresponding Asn cycle to the consensus sequence NX(S/T). The purified glycopeptides were further analyzed by MALDI-mass spectrometry, indicating the mass of carbohydrate moiety after subtracting peptide mass. Another strategy of carbohydrate analysis was performed by deglycosylation using *N*-glycanase digestion or hydrazinolysis. The sugar components were derivatized with 2-aminobenzamide and sodium borohydride (12). The derivatives were purified by reversed phase HPLC for analysis by mass spectrometry.

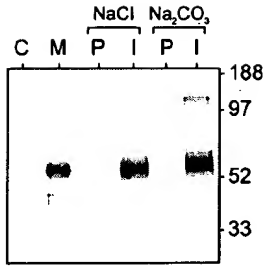



FIG. 1. **BACE is an integral membrane protein.** Immunoprecipitation of BACE from stably expressing 293 cells after overnight labeling is shown. *C*, cytosolic fraction; *M*, membrane fraction. Washing the membrane fraction with NaCl or $Na_2CO_3$ leads to the release of peripheral membrane proteins (*P*) into the wash phase, whereas integral membrane proteins (*I*) stay in the membrane.

## RESULTS

*BACE Is a Glycosylated Integral Membrane Protein That Is N-terminally Processed in the Golgi Apparatus*—Analysis of the BACE protein sequence suggests that BACE is a single transmembrane domain protein (7), and it has been shown that active enzyme can be released from membrane fractions after treatment with 0.2% Triton (9). We prepared cell lysates of 293 cells stably overexpressing BACE and separated cytoplasm and membrane fraction by ultracentrifugation. Immunoprecipitation with the BACE C-terminal antiserum (7) confirmed that BACE is present only in the membrane (*M*), but not in the cytoplasmic fraction (*C*) (Fig. 1). Washing the membrane fraction with 0.5 M sodium chloride or 0.1 M sodium carbonate, pH 11, does not release the protein into the wash phase (*P*), demonstrating that BACE is indeed an integral membrane protein (*I*). As noted before, mature BACE migrates on gels at ~70 kDa, a higher molecular mass than predicted from the amino acid sequence, suggesting that it may be glycosylated (7). When BACE is immunoprecipitated after 20 min labeling from the stable 293 line with the C-terminal antibody, an immature species running at ~60 kDa is detected (Fig. 2*A*, *lane B*). As expected, nontransfected control cells treated the same way do not show this band (*lane C*). If the immunoprecipitate is pretreated with *N*-glycosidase F (*lane BF*), the band runs at less than 50 kDa, indicating that the immature species is *N*-glycosylated. The same result is obtained with a second antibody raised to the propeptide region of BACE (Fig. 2*A*). This antibody does not show a band with non-transfected control cells (*lane C*) but recognizes the same 60-kDa *N*-glycosylated species as the C-terminal antibody, and the same molecular weight shift is observed upon *N*-glycosidase F treatment.

To analyze the turnover of BACE in the stable cell line, we performed a pulse-chase experiment in which the cells were labeled for 20 min and then chased in the absence of label. Cell lysates were prepared at the indicated times and immunoprecipitated with the C-terminal antibody (Fig. 2*B*). At time 0 immediately after labeling a strong 60-kDa band representing the immature *N*-glycosylated species is detectable. At 3 h this band has disappeared, and less than half of the original material is recovered as mature glycosylated 70-kDa form that is degraded slowly (*broken line*, $T_{1/2}$ >9 h, Fig. 2*D*). Thus, overexpressed BACE is glycosylated, and the immature *N*-glycosylated form is rapidly degraded. The immature *N*-glycosylated protein that escapes degradation is turned into the mature glycosylated form that is stable in 293 cells. We also performed pulse-chase experiments with the propeptide antibody (Fig. 2*C*). At time 0 the same 60-kDa band is detected as with the C-terminal antibody; however, by 2 h chase time most of

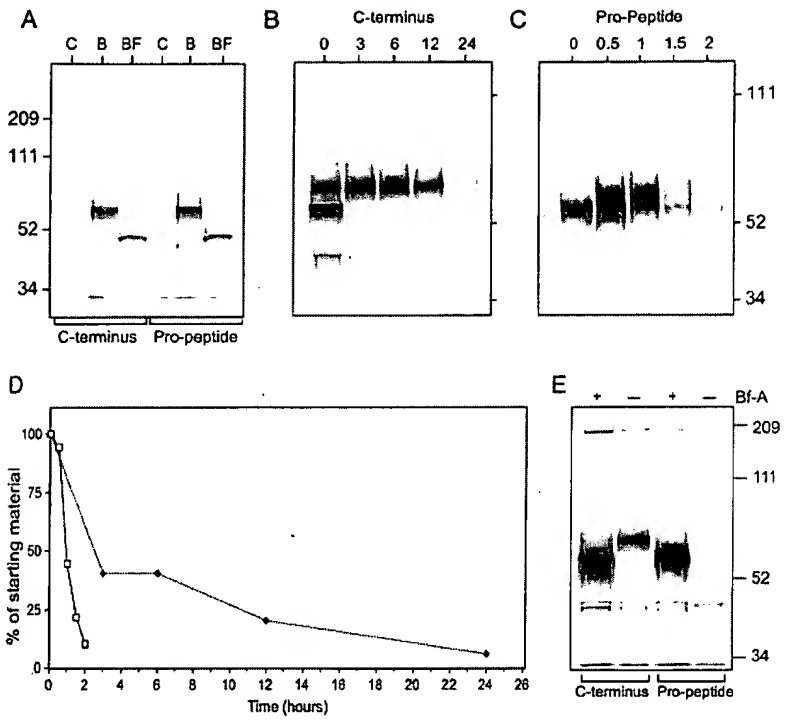



FIG. 2. **BACE processing and glycosylation.** *A*, immunoprecipitation of BACE from 293 cells after 20 min labeling using a C-terminal antiserum or a propeptide antiserum. *C*, nontransfected control cells; *B*, 293 cells stably expressing BACE; *BF*, samples from 293 cells stably expressing BACE treated with *N*-glycosidase F. *B*, immunoprecipitation of BACE from stably transfected 293 cells after 20 min labeling followed by the indicated chase times in hours using the C-terminal antiserum. *C*, immunoprecipitation of BACE from stably transfected 293 cells after brief labeling followed by the indicated chase times in hours using the propeptide antiserum. *D*, quantitation of the BACE signal by phosphorimaging. *Solid line*, propeptide signal; *broken line*, C-terminal signal. *E*, immunoprecipitation of BACE from cells chased for 3 h in the presence (+) or absence (−) of Brefeldin A.

the signal has disappeared (Fig. 2C, quantitation see *solid line* in Fig. 2*D*). Only a minor portion of the material is at a molecular mass higher than 60 kDa. These results indicate that the BACE protein undergoes constitutive N-terminal processing and that the N-terminal processing occurs in temporal proximity with the trimming/adding of carbohydrate residues of the immature form, *i.e.* in the Golgi apparatus. This finding was confirmed by a Brefeldin A treatment experiment (Fig. 2*E*). Cells chased for 3 h in the absence of Brefeldin A show only the mature 70-kDa protein that is detectable with the C-terminal antibody but not with the propeptide antibody. In contrast, cells chased with Brefeldin A for 3 h show the immature 60-kDa form that is detectable with both the C-terminal and the propeptide antibody. Thus, treatment of the cells with Brefeldin A blocks both N-terminal processing and further glycosylation of the immature 60-kDa form, indicating that propeptide cleavage happens in the Golgi apparatus.

*BACE-IgG Shows N-Glycosylation but Insignificant O-Glycosylation*—In order to characterize the posttranslational modifications of BACE in detail, it is necessary to purify a large quantity of the protein to homogeneity. We have previously described a soluble form of BACE that retains enzymatic activity but can be more easily purified than the transmembrane form. Because this fusion protein shows enzymatic activity, it is assumed that the structure of the BACE ectodomain is not compromised in a major way (7) and only with the fusion protein were we able to get sufficient material for biochemical characterization. This fusion protein has been termed BACE-IgG and contains the extracellular domain of β-secretase (residues 1–460) and the Fc portion (230 amino acids) of human γ-immunoglobulin as shown in Fig. 3*A*. Because the IgG portion of the fusion protein forms the homodimeric Fc piece, we expected to find a molecule of the structure $(BACE)_2$-$(IgG)_2$ in which the two IgG molecules are connected by intermolecular disulfide bonds. The fusion protein was expressed in human embryonic kidney 293 cells and purified from the conditioned media by protein A affinity chromatography, followed by gel filtration on Sephacryl S-300-HR. On non-reducing SDS-PAGE BACE-IgG showed a single band at approximately 116 kDa (Fig. 4, *lane 2*). An exact measure of protein mass was subsequently obtained by MALDI-mass spectrometry, revealing a single component with a molecular mass of 116 kDa (Fig. 5), consistent with the SDS-PAGE result. This molecular mass suggests the structure BACE-$(IgG)_2$ but not $(BACE)_2$-$(IgG)_2$ (see Fig. 3*B*). Consistent with the proposed structure BACE-$(IgG)_2$, SDS-PAGE after reducing treatment of purified BACE-IgG with β-mercaptoethanol shows the monomeric BACE-IgG fusion running at 90 kDa, as described previously (7), and the IgG piece running at approximately 30 kDa (Fig. 4, *lane 3*). Nonreducing SDS-PAGE after treatment with *N*-glycanase (*lane 6*), *O*-glycanase (*lane 8*), sialidase (*lane 9*), and sialidase + *O*-glycanase (*lane 10*) shows that BACE contains multiple *N*-glycosylation sites but insignificant *O*-glycosylation.

*N-terminal Processing of BACE-IgG*—Full-length BACE isolated from transfected cells (7) or from human brain (9) starts predominantly at position 46, suggesting efficient proprotein processing. Ten cycles of sequence analysis for the purified BACE-IgG showed multiple N-terminal sequences. Two sequences were derived from the N-terminal domain of BACE starting from residues 22 and 46, corresponding to sequences TQHGIRLPLR-(22–31) and ETDEEPEEPG-(46–55). We interpret the 22-form as the pro-form of the enzyme after cleavage of the signal peptide and the 46-form as the mature active species. The third sequence comes from the IgG portion, corresponding to AVTDKTHTXP-(461–470) for 10 residues. The ratio of the components was roughly 1:1 for BACE to IgG, as expected for BACE-$(IgG)_2$ (see Fig. 3*B*).

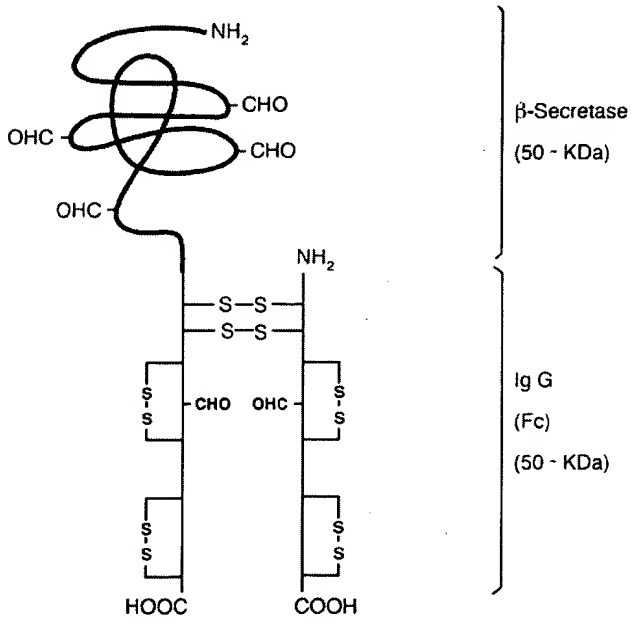
*Structural Analysis of BACE-IgG*—Purified BACE-IgG was examined for the presence of free sulfhydryl residues using FM labeling. The FM-labeled protein was digested with trypsin and endoproteinase Asp-N. The peptide mapping analysis (data not shown) indicated a few fluorescent-positive peaks. However, none of them gave N-terminal sequences. Thus, the FM-positive peaks might all be derived from the fluorescent reagent. This result suggests that all 6 cysteine residues in the BACE ectodomain form disulfide bonds. To analyze directly the disulfide bonds and the glycosylation sites, BACE-IgG was digested with trypsin and endoproteinase AspN. The double digestion

A

| | | | | | |
|---|---|---|---|---|---|
| 1 | MAQALPWLLL | WMGAGVLPAH | GTQHGIRLPL | RSGLGGAPLG | LRLPRETDEE |
| 51 | PEEPGRRGSF | VEMVDNLRGK | SGQGYYVEMT | VGSPPQTLNI | LVDTGSSNFA |
| 101 | VGAAPHPFLH | RYYQRQLSST | YRDLRKGVYV | PYTQGKWEGE | LGTDLVSIPH |
| 151 | GPNVTVRANI | AAITESDKFF | INGSNWEGIL | GLAYAEIARP | DDSLEPFFDS |
| 201 | LVKQTHVPNL | FSLQLCGAGF | PLNQSEVLAS | VGGSMIIGGI | DHSLYTGSLW |
| 251 | YTPIRREWYY | EVIIVRVEIN | GQDLKMDCKE | YNYDKSIVDS | GTTNLRLPKK |
| 301 | VFEAAVKSIK | AASSTEKFPD | GFWLGEQLVC | WQAGTTPWNI | FPVISLYLMG |
| 351 | EVTNQSFRIT | ILPQQYLRPV | EDVATSQDDC | YKFAISQSST | GTVMGAVIME |
| 401 | GFYVVFDRAR | KRIGFAVSAC | HVHDEFRTAA | VEGPFVTLDM | EDCGYNIPQT |
| 451 | DESTLMTIAY | [AVTDKTHTCP | PCPAPELLGG | PSVFLFPPKP | KDTLMISRTP |
| 501 | EVTCVVVDVS | HEDPEVKFNW | YVDGVEVHNA | KTKPREEQYN | STYRVVSVLT |
| 551 | VLHQDWLNGK | EYKCKVSNKA | LPAPIEKTIS | KAKGQPREPQ | VYTLPPSRDE |
| 601 | LTKNQVSLTC | LVKGFYPSDI | AVEWESNGQP | ENNYKTTPPV | LDSDGSFFLY |
| 651 | SKLTVDKSRW | QQGNVFSCSV | MHEALHNHYT | QKSLSLSPGK] | |

FIG. 3. *A*, sequence of BACE-IgG. The Fc sequence (461–690, in *brackets*) of human IgG1 was attached to the ectodomain of BACE. *B*, schematic model of the BACE-IgG fusion protein. The BACE-$(IgG)_2$ structure determined in this study is shown. Estimated molecular masses (kDa) are based on the protein sequence, not including carbohydrate moieties.

B



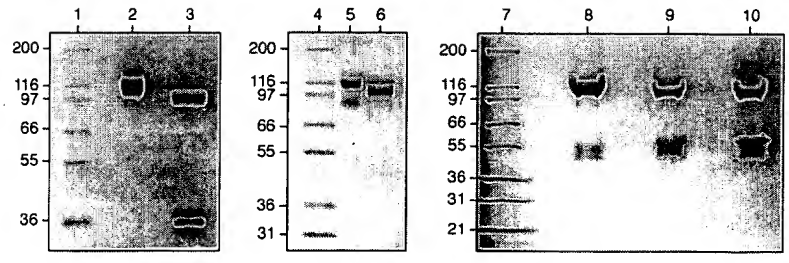




FIG. 4. **SDS-PAGE of purified BACE-IgG.** The sample was loaded onto a nonreducing gel (4–20%) with SDS buffer. *Lanes 1, 4,* and *7*, molecular weight markers; *lanes 2* and 5 BACE-IgG untreated; *lane 3*, after reduction with β-mercaptoethanol; *lane 6*, *N*-glycanase-treated sample; *lane 8*, *O*-glycanase-treated sample; *lane 9*, sialidase-treated sample; *lane 10*, sialidase + *O*-glycanase-treated sample. Bands in *lane 8–10* around the 55-kDa marker were from *O*-glycanase or sialidase.

was performed to obtain the Cys-containing peptides or glycopeptides. The peptide map (data not shown) demonstrated significant resistance against the serine protease or endoproteinase Asp-N, so peptide recovery was insufficient. Nevertheless, several peptides gave useful information for elucidating the structure. The results are summarized in Table I. Peptide TD22.8 gave two sequences, D*X*K-(277–279) and D*X*GYNIPQT-(442–450), where *X* is to be a cysteine residue according to the amino acid sequence (Fig. 3*A*). Mass spectrometry supported the conclusion that the peptides were linked between these cysteines. Peptide TD27.5a showed a similar, but C-terminally extended sequence D*X*KEYNY-(277–283).

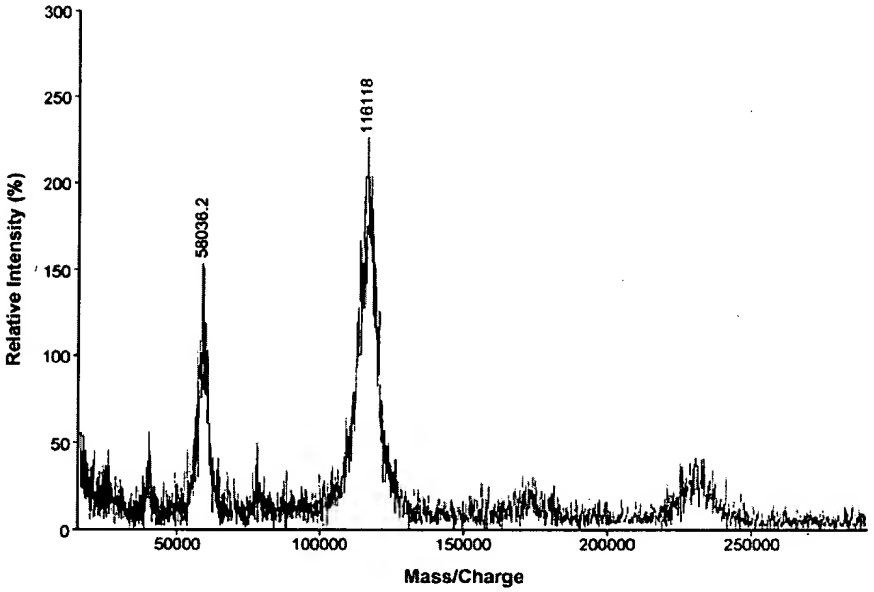



FIG. 5. **MALDI-mass spectrometry of the purified enzyme.** The BACE-IgG protein sample was loaded onto a slide with the matrix sinapinic acid. Protein mass was analyzed using a Voyager mass spectrometer as described under "Experimental Procedures." The mass at 116 kDa represents the singly charged ion and the mass at 58 kDa the doubly charged ion.

TABLE I
*Sequences of Cys-containing peptides from trypsin-endoproteinase Asp-N double digestion of BACE-IgG*

| Peptide | Sequence (residue no.)[a] | Observed mass | Calculated mass ($MH^+$) | Disulfide bond |
|---|---|---|---|---|
| TD 22.8 | DCK-(277–279)<br>DCGYNIPQT-(442–450) | 1375 | 1376.5 | $Cys^{278}$–$Cys^{443}$ |
| TD27.5a | DCKEYNY-(277–283)<br>DCGYNIPQT-(442–450) | 1942 | 1946.1 | $Cys^{278}$–$Cys^{443}$ |
| TD27.5b | TPEVTCVVVD-(499–508)<br>CK-(564–565) | 1194 | 1197.4 | $Cys^{504}$–$Cys^{564}$ |

[a] Cysteine residues were not detected by sequence analysis, but they are derived from the protein sequence (see Fig. 3*A*).

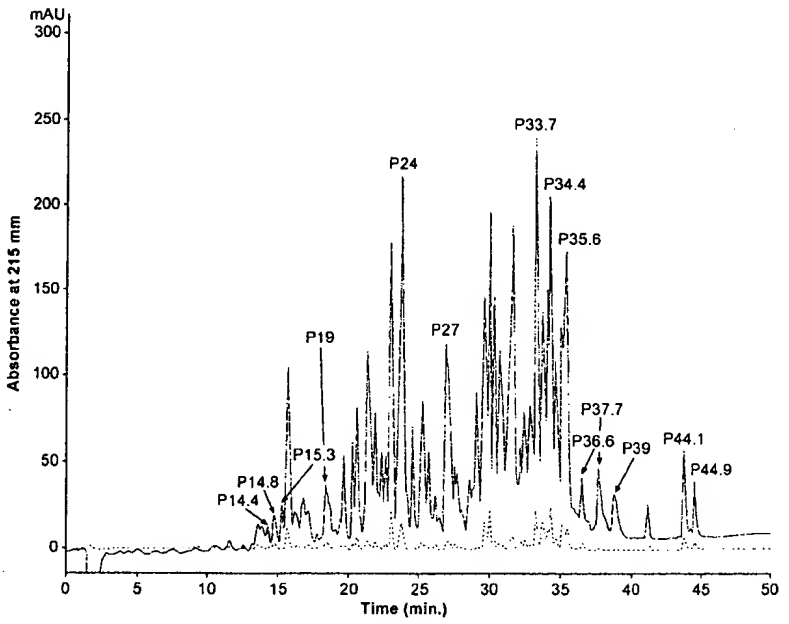



FIG. 6. **HPLC map of pepsin-generated peptides from BACE-IgG.** The digested sample was subjected to reversed phase HPLC as described in text. The peptides were detected by absorbances at 215 nm (*solid line*) and 280 nm (*dotted line*).

Mass spectrometry of both peptides confirmed the disulfide linkages as indicated with masses of 1376.5 and 1946.1, respectively. From these results we assign the first disulfide linkage as $Cys^{278}$–$Cys^{443}$. Finally, peptide TD27.5b consisting of the two peptides, *X*K-(564–565) and TPEVT*X*VVVD-(499–508), indicates the presence of $Cys^{504}$–$Cys^{564}$ in the Fc region. Since we could not obtain sufficient information to determine all disulfide bonds from the TD-digested peptides alone, the protein was digested with pepsin under acidic conditions. The pepsin-generated peptide map is shown in Fig. 6. Sequence analysis and mass spectrometry revealed the key peptides for determining disulfide linkages and *N*-glycosylation sites, and the analyzed disulfide bond containing peptides are shown in Table II. Peptide P33.7 contained two sequences LKMD*X*KEY-(274–281) and DMED*X*GYNIPQT-(439–450). Mass spectrometry confirmed this assignment although the observed mass was

TABLE II
*Sequences of pepsin-generated Cys peptides of BACE-IgG*

| Peptide | Sequence (residue no.) | Observed mass | Calculated mass ($MH^+$) | Disulfide bonds |
|---|---|---|---|---|
| P30.5 | AVTDKTHTCPPCPAPEL(LG)-(461–477/461–479)[a] | 3730 | 3734.4 | $Cys^{469}$–$Cys^{472}$ |
| P33.7 | LKMDCKEY-(274–281)<br>DMEDCGYNIPQT-(439–450) | 2552 | 2532.8 | $Cys^{278}$–$Cys^{443}$ |
| P34.6 | TCLVKGFYPSD-(609–619)<br>SCSVMHEALHNHYTQKS-(667–683) | 3902 | 3903.1 | $Cys^{610}$–$Cys^{668}$ |
| P39.0 | FSLQLCGAGFPLNQSEVL-(211–228)<br>AVSACHVHDEF-(416–426) | [b] | 3139.5 | $Cys^{216}$–$Cys^{420}$ |
| P44.9 | LVCWQAGTTPWNIF-(328–341)<br>VATSQDDCYKF-(373–383) | 2915 | 2914.3 | $Cys^{330}$–$Cys^{380}$ |

[a] Two similar peptides (461–477 and 461–479) were cross-linked. We did not determine which of the cysteines form the disulfide bonds.
[b] Due *to* *N*-glycosylation, the sample did not show the expected mass (see Table III).

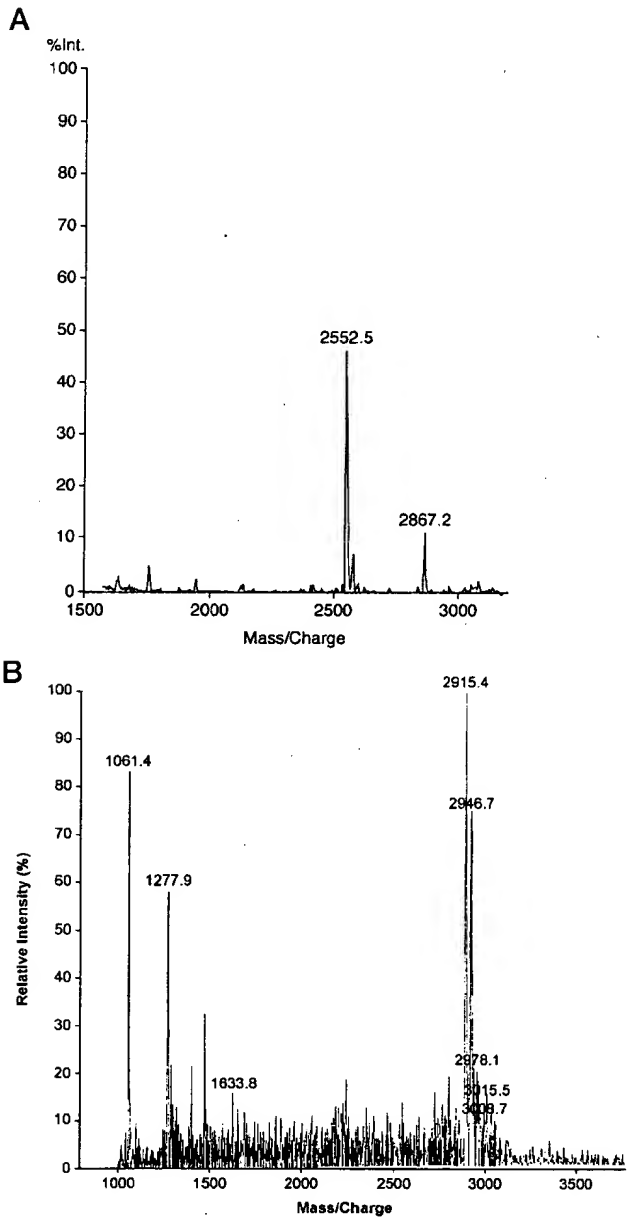



FIG. 7. **Mass spectrometry of Cys-containing peptides.** *A*, mass spectrum of P33.7. *B*, mass spectrum of P44.9. Peptide mass was determined by MALDI-mass spectrometer using Kratos IV. The sample was loaded onto a slide with 4-HCCA as matrix.

slightly higher than the expected probably due to oxidation of a methionine residue (Fig. 7*A*). This disulfide bond was already assigned by peptide TD22.8 (see above). Peptide P37.7 and P39 showed two sequences, FSLQL*X*GAGFPLNQSEVL-(211–228) and AVSA*X*HVHDEF-(416–426), where *X* indicates the cysteine residue. This peptide permitted us to determine $Cys^{216}$–$Cys^{420}$. Asn at residue 223 was not detected by sequence analysis because of the *N*-glycosylation. The difference between the peptides P37.7 and P39 may be due to carbohydrate heterogeneity. Mass spectrometry of the peptide was not successful due to glycosylation. The third disulfide linkage $Cys^{330}$–$Cys^{380}$ was determined by analysis of peptide P44.9, containing two sequences, LV*X*WQAGTTPWNIF-(328–341) and VATSQDD*X*Y-KF-(373–383). The observed mass of 2915 from peptide P44.9 was consistent with the predicted mass 2914.3 within the experimental errors (Fig. 7*B*). Another disulfide of the Fc portion was determined to be $Cys^{610}$–$Cys^{668}$ from peptide P34.6 (see Table II). Finally, the dimerized peptide P30.5 demonstrates the intermolecular linkages between $Cys^{469}$ and $Cys^{472}$ in the Fc portion.

*N-Glycosylation Sites*—We have analyzed the glycopeptides for the identification of carbohydrate attachments (Table III). Four *N*-glycosylation sites ($Asn^{153}$, $Asn^{172}$, $Asn^{223}$, and $Asn^{354}$) from BACE and one site ($Asn^{540}$) from IgG are predicted according to the consensus sequence, N*X*(S/T). After sequence analysis of all peptic peptides, we found that the four potential *N*-glycosylation sites of BACE are indeed occupied by carbohydrate moieties. The fact that glycopeptides with the same amino acid sequence were separately eluted on HPLC suggests that these *N*-glycosylation sites may have carbohydrate heterogeneities. For example, the peptide FINGSN-(170–175) containing $Asn^{172}$ is separated into several peaks, P14.4, P14.9, and P15.3, respectively (Fig. 6). Moreover, mass spectrometry of a single HPLC peak, *e.g.* P27, gave several mass units, 3158.5, 3320.7, 3524.1, and 3686.2, respectively. According to the sequence analysis the glycopeptide P27 has the sequence VSIPHGPNVTVRA-(146–158) (mass = 1347.5) and *N*-glycans of this peptide should have 1811.0, 1973.2, 2176.6, and 2338.7 mass units, respectively. Thus, even considering experimental errors, our mass data predict multiple carbohydrate structures. Sequence and mass spectral analyses of the glycopeptides are listed in Table III. Due to the complexity of the problem, determination of the exact carbohydrate structure is still in progress, but the mass spectral fragmentation suggests that the predicted carbohydrates may have high hexose units, leading to the observed structural heterogeneity.

## DISCUSSION

This study provides the first characterization of the recently identified β-secretase protein BACE at the biochemical level. We began our analysis by addressing properties of the intact form of BACE that contains the predicted transmembrane domain, and we confirmed that the BACE protein is an integral membrane protein (7, 9). Analysis of the turnover of BACE in overexpressing 293 cells demonstrated that BACE is constitutively processed to a mature form lacking the propeptide re-

TABLE III
*Glycopeptides from BACE-IgG*

| Peptide | Sequence (residue no.) | Sites[a] | Observed mass (Calculated) |
|---|---|---|---|
| P14.8 | **FINGSN**-(170–175) | $Asn^{172}$ | ND (651.7) |
| P19.0 | EVT**NQSF**-(351–357) | $Asn^{354}$ | 3511.9 (824.8) |
| | | | 3674.0 |
| P24.0 | YVDGVEVHNAKTKPREEQ**YNST**-(521–542) | $Asn^{540}$ | 4148.2 (2565.7) |
| | | | 4309.8 |
| | | | 4471.6 |
| P27.0 | VSIPHGP**NVT**VRA-(146–158) | $Asn^{153}$ | 3158.5 (1347.5) |
| | | | 3320.7 |
| | | | 3524.1 |
| | | | 3686.2 |
| P39.0 | FSLQLCGAGFPL**NQS**EVL-(211–228) | $Asn^{223}$ | 4120 (3139.5) |
| | AVSACHVHDEF (416–426) | | 4552 |

[a] *N*-Glycosylation sites were determined by no detection of PTH-Asn at the corresponding cycle. Boldface letters show consensus sequences N*X*(S/T) for *N*-glycosylation. The peptide P39.0 contained two peptides cross-linking through a disulfide bond.

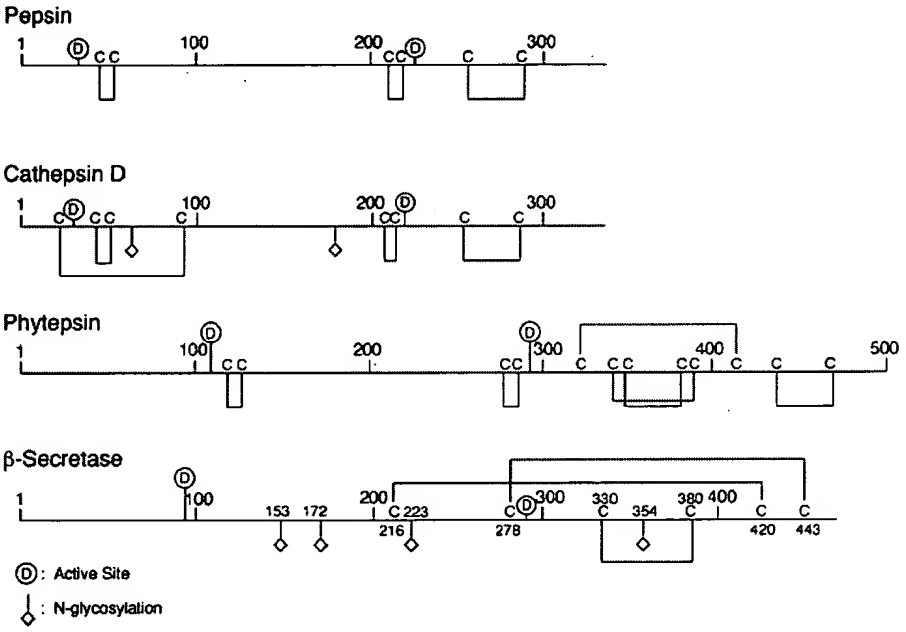



FIG. 8. **Comparison of disulfide motif and *N*-glycosylation sites in aspartic proteases.** The ectodomain of the β-secretase enzyme BACE has full enzyme activity and contains the three intramolecular disulfide bonds as determined here. In comparison with other aspartic proteases like pepsin and cathepsin D, BACE contains a different disulfide connectivity. Active site Asp (*D*) residues are *circled*.

gion. Apparently, this processing is quite efficient, even under overexpression conditions, as has been reported before (9). At least in the 293 cells tested here processing of BACE does not appear to limit β-secretase activity. The immature BACE protein is rapidly turned over, and less than half of the initial material is recovered as mature protein. We do not know at this point whether the massive loss of immature protein is an overexpression artifact or whether a major proportion of immature BACE is degraded under low level expression conditions as well. Once processed, BACE is quite stable even under overexpression conditions. Our results show that BACE is glycosylated. The findings that there is almost no fully glycosylated BACE which still contains the propeptide epitope and that Brefeldin A treatment blocks processing indicate that the cleavage of the propeptide happens in the Golgi apparatus. The nature of the propeptide processing enzyme is currently under investigation, but an autocatalytic mechanism, as reported for pepsin (13), seems unlikely, if one considers the sequence specificity of BACE (7).

To analyze the biochemistry of BACE in more detail, we made use of the previously described BACE-IgG construct (7) containing the entire ectodomain of BACE, which can be purified much more conveniently than the transmembrane form. Because this form of the enzyme is active and maintains the sequence specificity of β-secretase (7), it appears justified to study structural features of BACE using this soluble form. Sequencing of BACE-IgG confirms the $Glu^{46}$ start previously described for the transmembrane form (7, 9) and also identifies a species starting at $Thr^{23}$ that has the signal peptide cleaved off, but still contains the propeptide. We observed much lower amounts of this form when we analyzed membrane-bound BACE, suggesting that the propeptide cleavage of BACE-IgG is not quite as efficient as that of BACE. Whether this is due to different transport kinetics or other differences between the two forms is currently not known.

The ectodomain of BACE contains six cysteines. According to the SH-labeling experiments it does not contain any free cysteines, but they all form disulfide bonds. Within BACE we did not detect dimeric forms caused by covalent intermolecular bonds, but instead we demonstrated that all three disulfide bonds are intramolecular linkages. Since BACE is clearly a member of the pepsin family (7), one might expect that it could have a structure similar to other aspartic proteases including pepsin, cathepsin D or E, and human immunodeficiency virus proteases (for review see Ref. 13). However, here we show that it has no significant homology with other pepsin family members in the disulfide structure. As shown in Fig. 8, only phytepsin, a plant aspartic protease (14), showed partial similarity with β-secretase in the big loops of the C-terminal domain. These structural differences may affect substrate specificity of the enzymes. Obviously, a detailed discussion of the structure function-relationship for β-secretase will require x-ray crystal-

lographic studies. Understanding this prime target for the treatment of Alzheimer's disease at the atomic level may turn out to be crucial for drug development.

REFERENCES


1. Alzheimer, A. (1907) *Centralbl. Nervenheilk. Psychiatr.* **30,** 177–179
2. Glenner, G. G., and Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* **120,** 885–890
3. Younkin, S. G. (1998) *J. Physiol. (Paris)* **92,** 289–292
4. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K.-H., Multhaup, G., Beyreuther, K., and Muller-Hill, B. (1987) *Nature* **325,** 733–736
5. Haass, C., and Selkoe, D. J. (1993) *Cell* **75,** 1039–1042
6. Citron, M., Teplow, D. B., and Selkoe, D. J. (1995) *Neuron* **14,** 661–670
7. Vassar, R., Bennett, B. D., Babu-Khan, S., Kahn, S., Mendiaz, E. A., Denis, P., Teplow, D. B., Ross, S., Amarante, P., Loeloff, R., Luo, Y., Fisher, S., Fuller, J., Edenson, S., Lile, J., Jarosinski, M. A., Biere, A. L., Curran, E., Burgess, T., Louis, J.-C., Collins, F., Treanor, J., Rogers, G., and Citron, M. (1999) *Science* **286,** 735–741
8. Hussain, I., Powell, D., Howlett, D. R., Tew, D. G., Meek, T. D., Chapman, C., Gloger, I. S., Murphy, K. E., Southan, C. D., Ryan, D. M., Smith, T. S., Simmons, D. L., Walsh, F. S., Dingwall, C., and Christie, G. (1999) *Mol. Cell. Neurosci.* **14,** 419–427
9. Sinha, S., Anderson, J. P., Barbour, R., Basi, G. S., Caccavello, R., Davis, D., Doan, M., Dovey, H. F., Frigon, N., Hong, J., Jacobson-Croak, K., Jewett, N., Keim, P., Knops, J., Lieberburg, I., Power, M., Tan, H., Tatsuno, G., Tung, J., Schenk, D., Seubert, P., Suomensaari, S. M., Wang, S., Walker, D., Zhao, J., Mc Conlogue, L., and Varghese, J. (1999) *Nature* **402,** 537–540
10. Yan, R., Bienkowski, M. J., Shuck, M. E., Miao, H., Tory, M. C., Pauley, A. M., Brashier, J. R., Stratman, N. C., Mathews, W. R., Buhl, A. E., Carter, D. B., Tomasselli, A. G., Parodi, L. A., Heinrikson, R. L., and Gurney, M. E. (1999) *Nature* **402,** 533–537
11. Haass, C., Schlossmacher, M. G., Hung, A. Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B. L., Lieberburg, I., Koo, E. H., Schenk, D., Teplow, D. B., and Selkoe, D. J. (1992) *Nature* **359,** 322–325
12. Anumula, K. R., and Dhume, S. T. (1998) *Glycobiology* **8,** 685–694
13. Davies, D. R. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 189–215
14. Kervinen, J., Tobin, G. J., Costa, J., Waugh, D. S., Wlodawer, A., and Zdanov, A. (1999) *EMBO J.* **18,** 3947–3955